A Statistical STT-RAM Retention Model for Fast Memory Subsystem Designs

Zihao Liu[1], Wujie Wen[1], Lei Jiang[2], Yier Jin[3] and Gang Quan[1]

[1]Florida International University, ,Miami, FL
[2]Indiana University Bloomington, Indiana
[3] University of Central Florida, FL

Presented by Wujie Wen

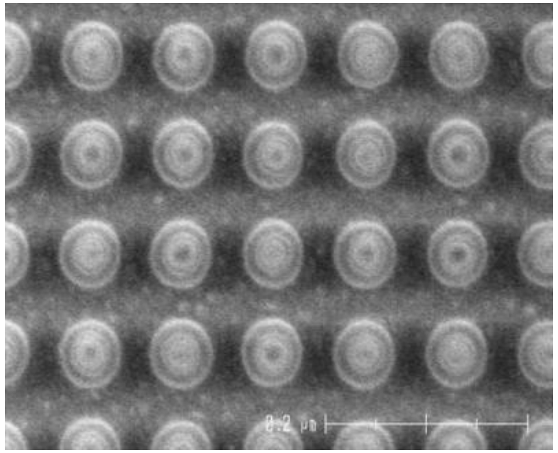
Memory Technologies Comparison

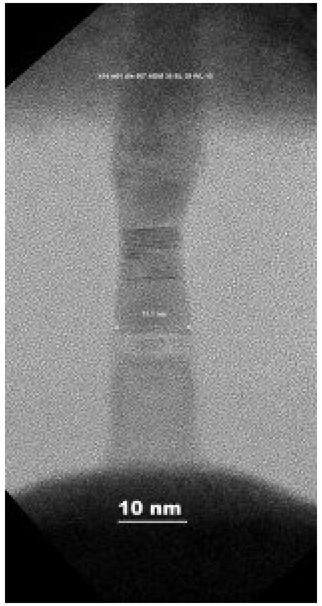
	SRAM	DRAM	NAND FLASH	STT-RAM	PCRAM	ReRAM
Data Retention	N	4 ms	10y	>10 y	>10 y	>10 y
Memory Cell (F^2)	120-140	7-9	4	8	4	<1
Read Time	0.2 ns	2 ns	0.1 ms	5-10 ns	12 ns	5-10 ns
Write/Erase Time	70 ps	1 ns	1/0.1ms	<10 ns	<50 ns	<10 ns
Number of Rewrites	10^{16}	10^{16}	10^{5}	10^{15}	10^{8}	10^{15}
Power Consumption Read/Write	Low	Low	High	Low	Low	Low
Power Consumption other than R/W	Leakage Current	Refresh Power	None	None	None	None

Source: ITRS ERD workshop presentation by Prof. Y. Chen

Recent MRAM News

1. Toshiba and Hynix prototype a 4 Gb STT-MRAM(Dec 20, 2016)
2. Samsung demonstrates a 8Mb embedded pMTJ STT-MRAM device(Dec 12, 2016)
3. IBM demonstrated 11nm STT-MRAM junction, says "time for STT-MRAM is now "(Jul 08, 2016)





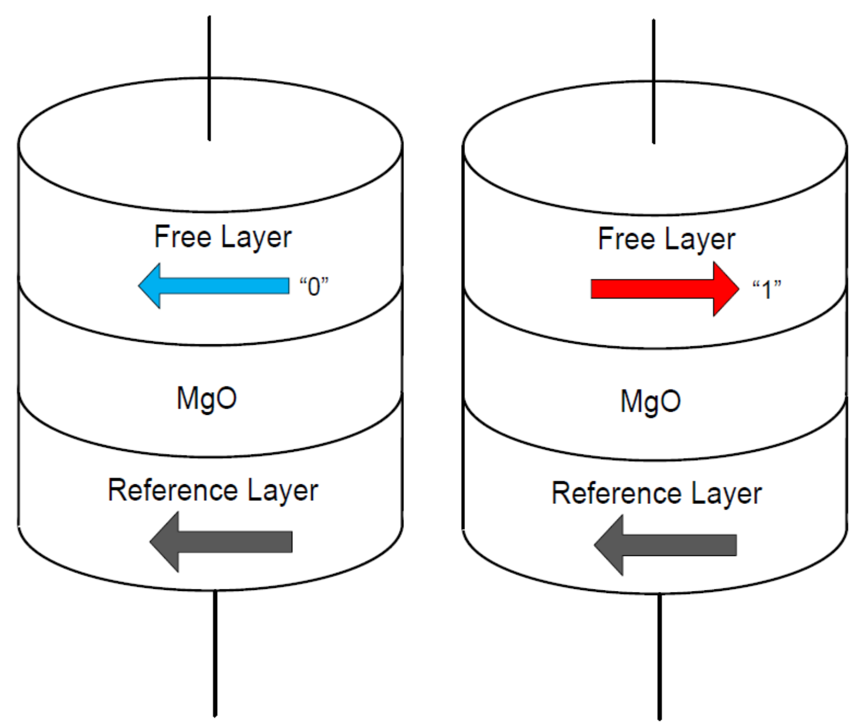
Outline

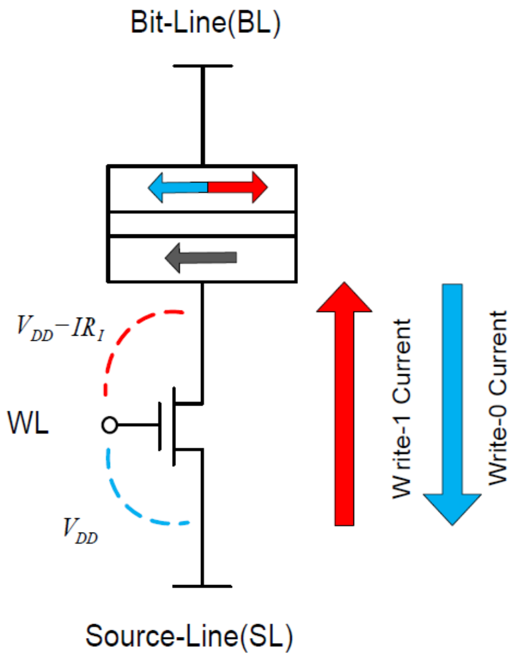
- Introduction

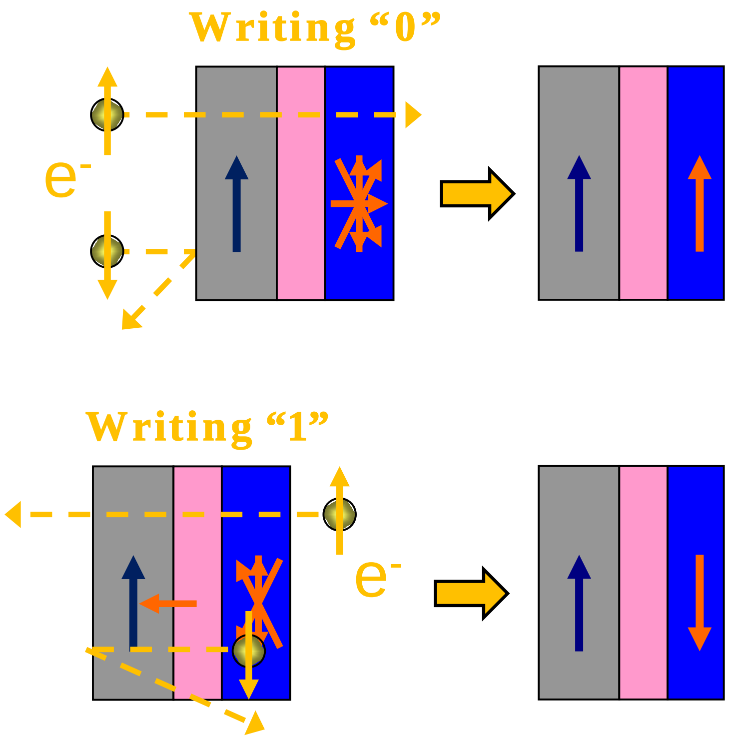
  ---STT-RAM Basic

  ---Retention Failure
- Modeling and Validation
- Methodology and Simulation
- Results
- conclusions

STT-RAM Basic







- Data '0'
  The magnetization directions of two layers are **parallel (low resistance)**
- Data '1'
  The magnetization directions of two layers are **anti-parallel (high resistance)**

Retention Failure

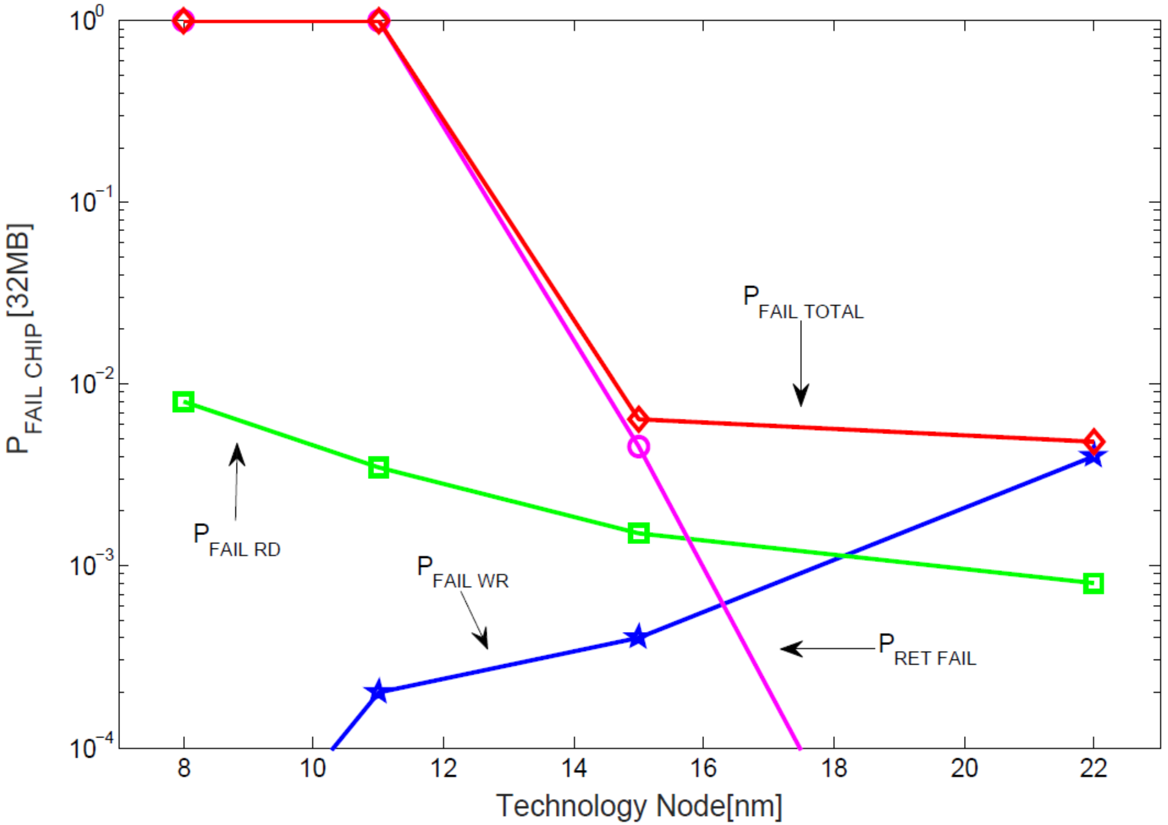
What is retention failure?

The retention failure of STT-RAM is a phenomenon where an idle cell flips without applying any intentional excitation source.

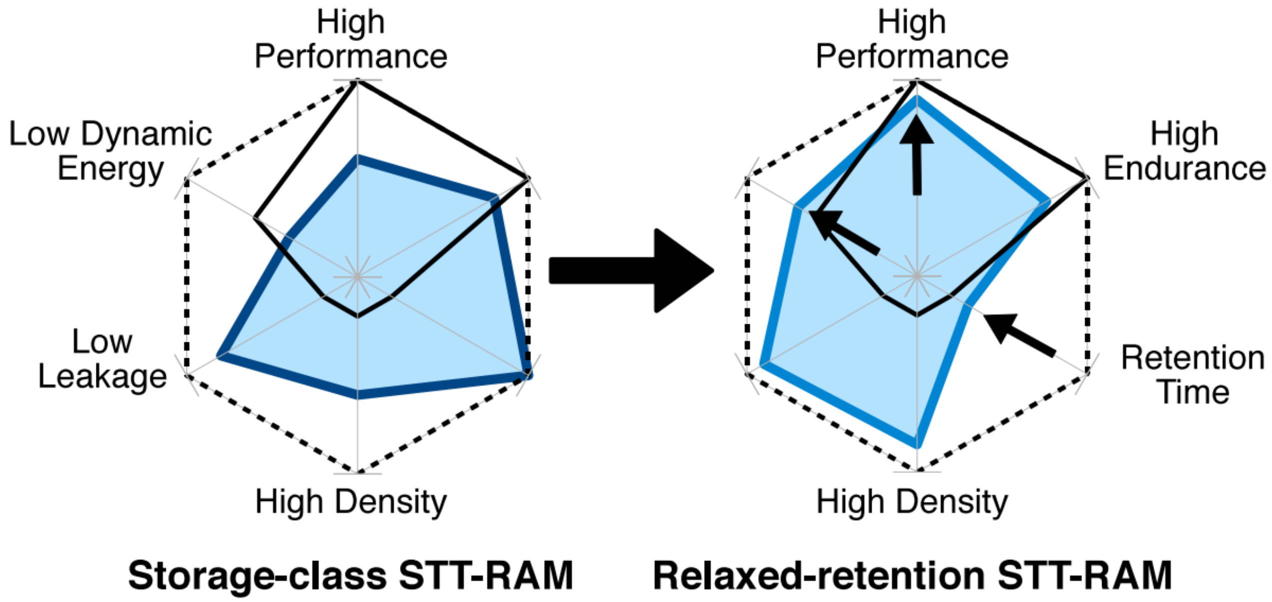
Two major reasons deteriorate retention failure:

- Process technology scales;
- Intentional relaxing STT-RAM retention to facilitate the fast and energy-efficient design.

Two major reasons cause retention failure



Source: Intel® Technology Journal | Volume 17, Issue 1, 2013



Source: Relaxing Non-Volatility for Fast and Energy-Efficient STT RAM Caches(C.W.Smullen, 2011)

The state of the art retention model

Previous works estimate the STT-RAM retention time using the following equation:

$$\ln(t) = \frac{A * t_m * H_k * M_s}{2k_b * T}$$

Three limitations of this method:

- The full **statistical information** of the retention time is not available;
- The **device and environmental variations** are not taken into considerations;
- This equation is derived by assuming an **excitation source–spin current is applied**.

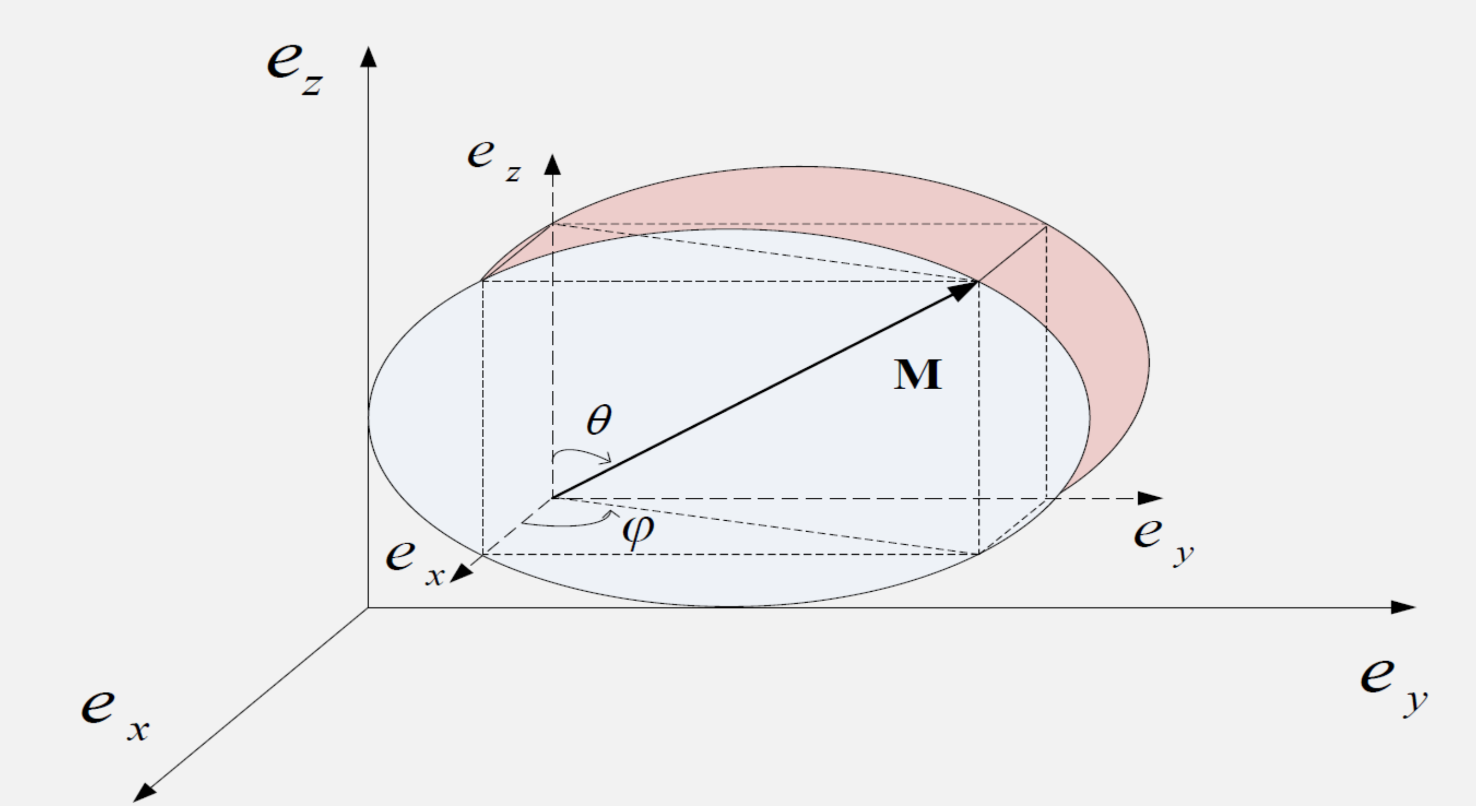
Outline

- Introduction
- Modeling and Validation

  --- Modeling

  --- Model Validation
- Methodology and Simulation
- Results
- conclusions

Modeling



θ --- the angle of vector M and z axis.

ϕ --- the angle of plane of vector M and z axis and x axis.

θ and ϕ describe the dynamic behavior of **free layer magnetization (M)** at any time.

- **θ = 0 for low resistant state.**
- **θ = π for high resistant state.**

Energy modeling

Three potential energy included in our model: $U(\theta,\varphi) = U_K + U_p + U_l$

- **The uniaxial anisotropy energy**;

$$U_K = K\sin^2\theta$$

- **The easy-plane anisotropy**;

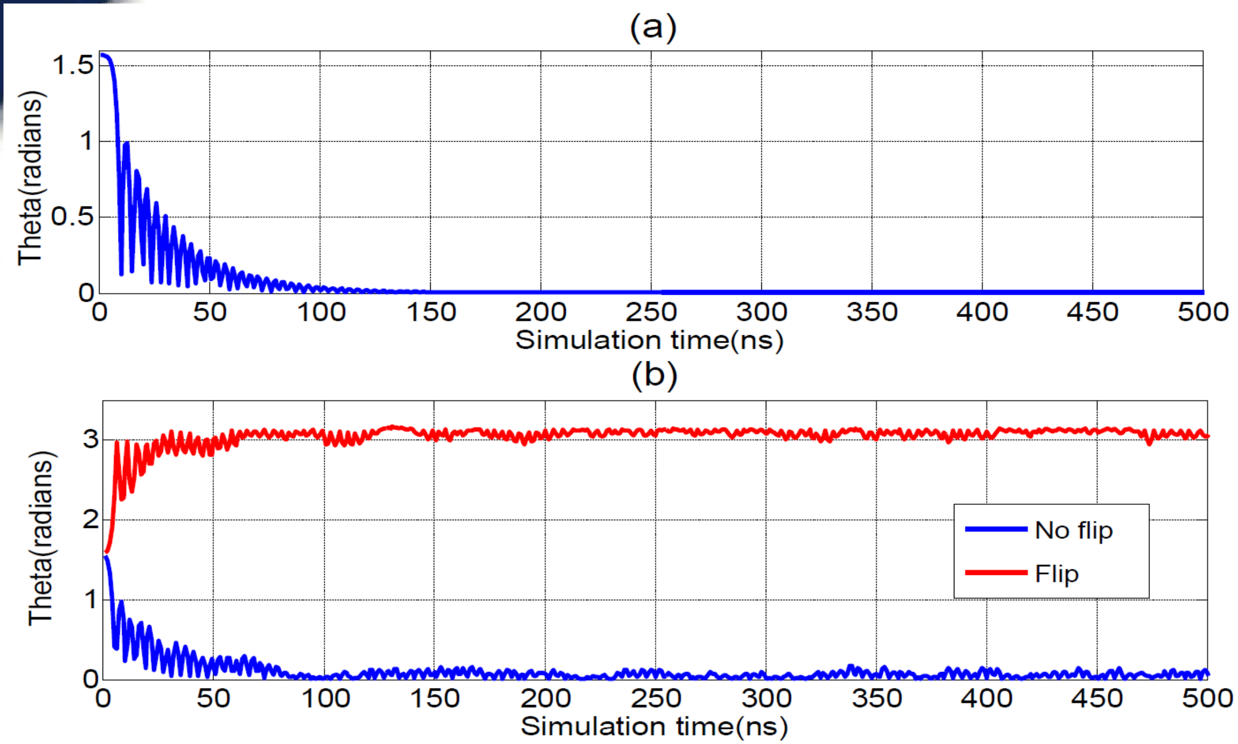
$$U_p = K_P(\sin^2\theta\cos^2\varphi - 1)$$

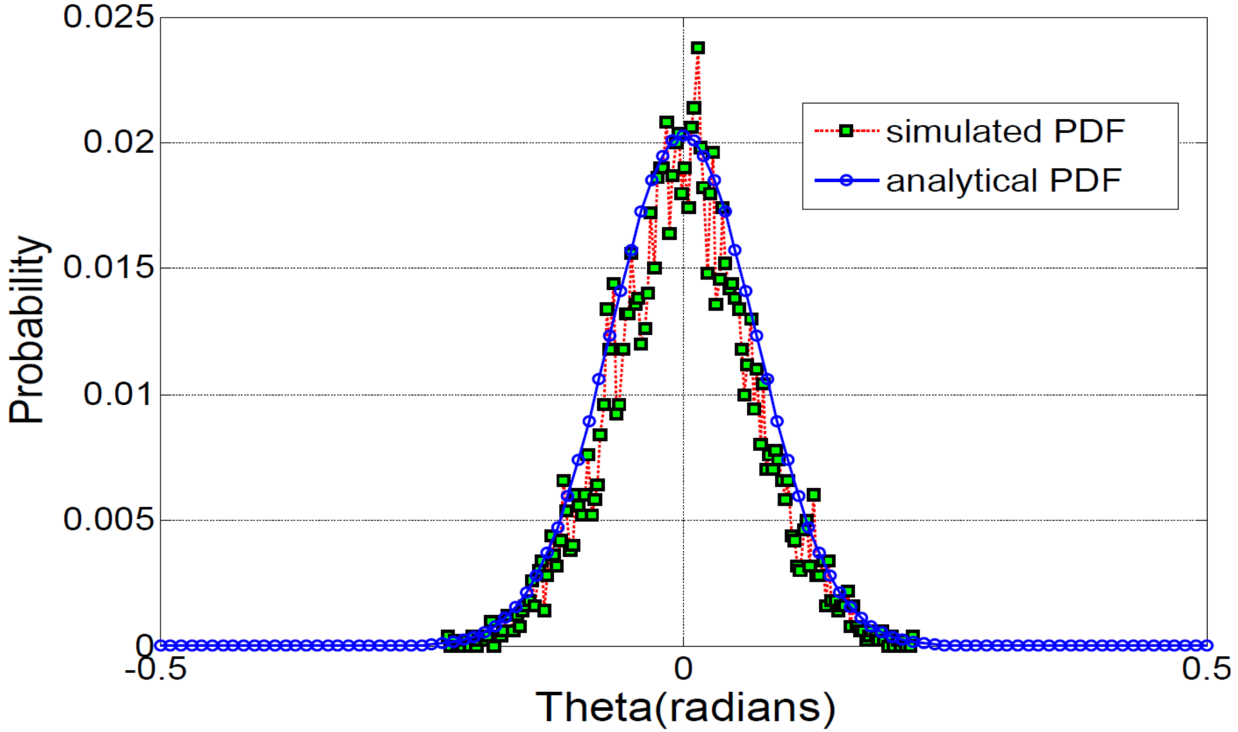
- **The thermal field energy**.

$$h_{l,i} = h_l \mathbf{X}_i(t) = \sqrt{\frac{\alpha}{\alpha^2+1}\frac{2k_bT}{\gamma\mu_0\Delta t\Delta v M_s}}\mathbf{X}_i(t)\ (i = x,y,z)$$

The aforementioned equation can be solved by **Landau–Lifshitz–Gilbert equation (LLG)** to capture the **free layer magnetization dynamic motion**.

Model Validation





The validation has two parts:

- **The dynamic behavior of θ with/without thermal field energy**;
- **The simulated PDF of θ compare to analytical results**.

Outline

- Introduction
- Modeling and Validation
- Methodology and Simulation

--- Methodology ("Create-Transfer-Recover-Calculate" method)

--- Simulation results validations

- Results
- conclusions

Methodology

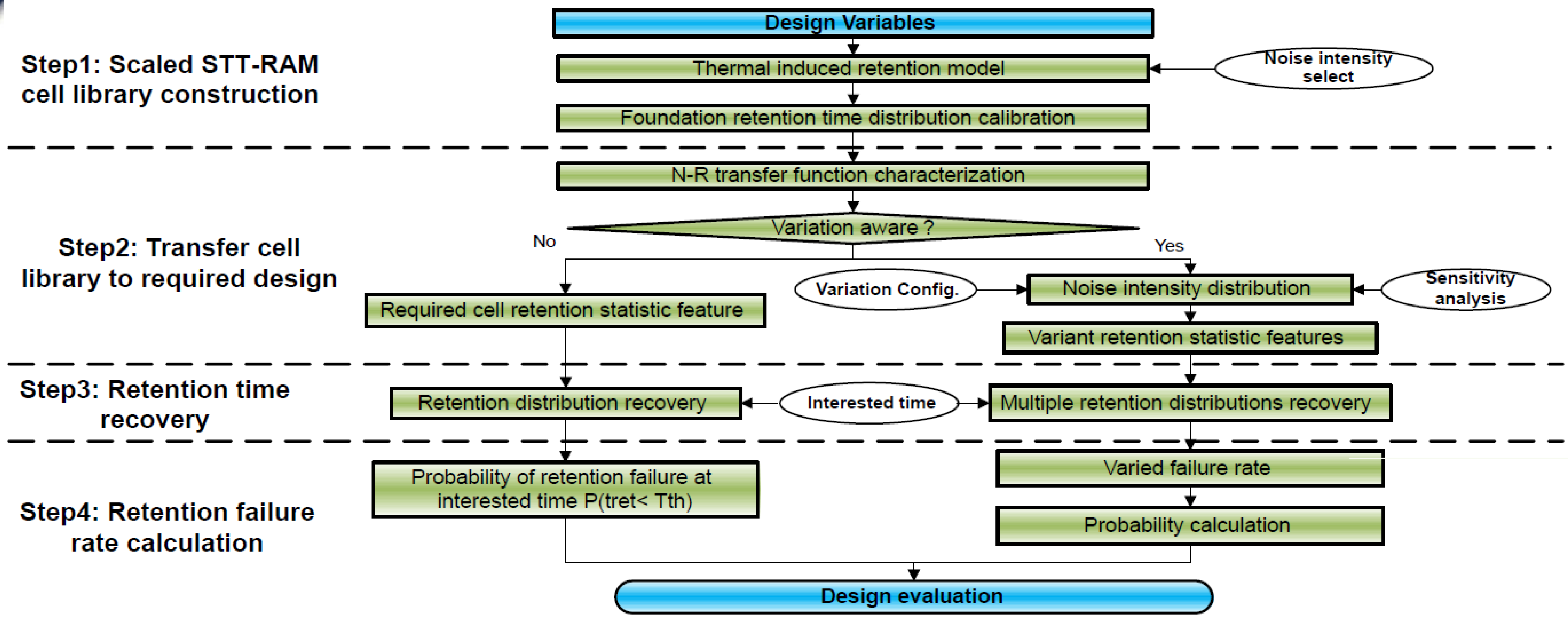
Retention time measurement is costly:

- **Retention failure is rarely happens** even in a sufficient long time window**;**
- **Lots of factors can affect** the stochastic retention behavior. (Volume, temperature etc.)
- **Parameters variations** make this measurement more complex.

Our proposed "Create-Transfer-Recover-Calculate" semi-analytical method:

- ***Create*** a "highly scaled STT-RAM cell" library for the fast retention time calibration;
- ***Transfer*** the statistical information from the library to any "desired STT-RAM cell";
- ***Recover*** the retention time distributions of the "desired STT-RAM cell";
- ***Calculate*** the retention failure rate within a given time for the normal mode or variation-aware mode.

Design flow



Design Variables
Step1: Scaled STT-RAM cell library construction
Thermal induced retention model
Noise intensity select
Foundation retention time distribution calibration
N-R transfer function characterization
Variation aware ?
No
Yes
Step2: Transfer cell library to required design
Required cell retention statistic feature
Variation Config.
Noise intensity distribution
Sensitivity analysis
Variant retention statistic features
Step3: Retention time recovery
Retention distribution recovery
Interested time
Multiple retention distributions recovery
Step4: Retention failure rate calculation
Probability of retention failure at interested time P(tret< Tth)
Varied failure rate
Probability calculation
Design evaluation

Step1: Create Distribution Model in Library

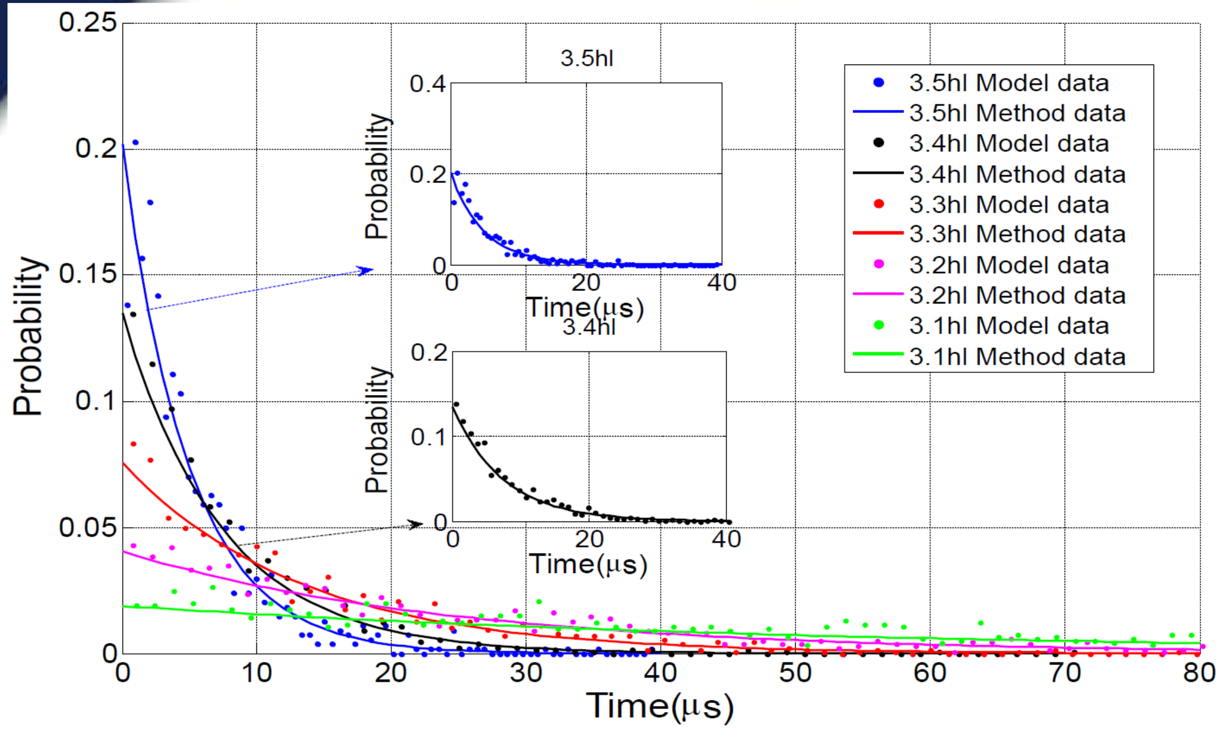
MAGNETIC SIMULATION RESULTS FOR SCALED CELLS

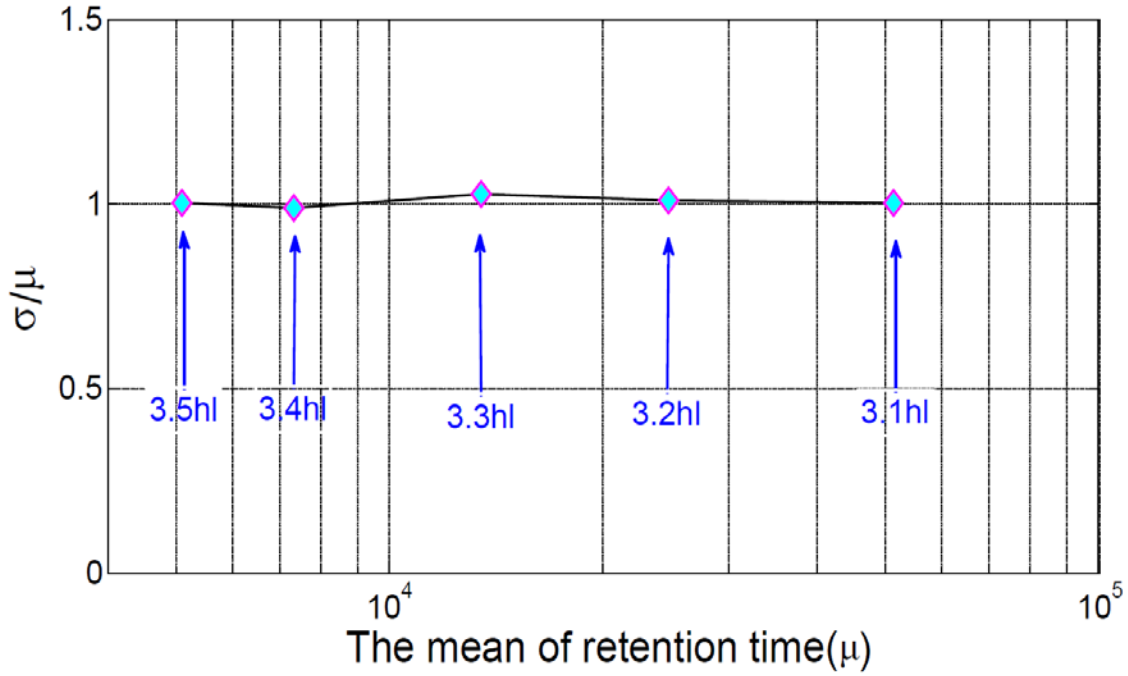
β	$T_{th}(ns)$	$P_f(T_{th})$	$\log\frac{-T_{th}}{\log(1-P_f)}$	$log(\mu)$
3.1	10^4	0.168	10.8834	N/A
	$5*10^4$	0.603	10.895	N/A
	10^5	0.857	10.847	N/A
	$2*10^6$	1	N/A	10.8733
3.5	10^4	0.874	8.472	N/A
	$2*10^4$	0.981	8.516	N/A
	$3*10^4$	0.996	8.491	N/A
	$4*10^4$	1	N/A	8.511

The **simulation results (last column)** of retention failure rate is always equal to **analytical results (the fourth column)**.

Consequently, the **retention failure rate** P_f for **a given time** T_{th} can be easily derived as:

$$P_f = P\left(t_{ret} \leq T_{th}\right) = 1 - e^{-T_{th}/\mu}$$



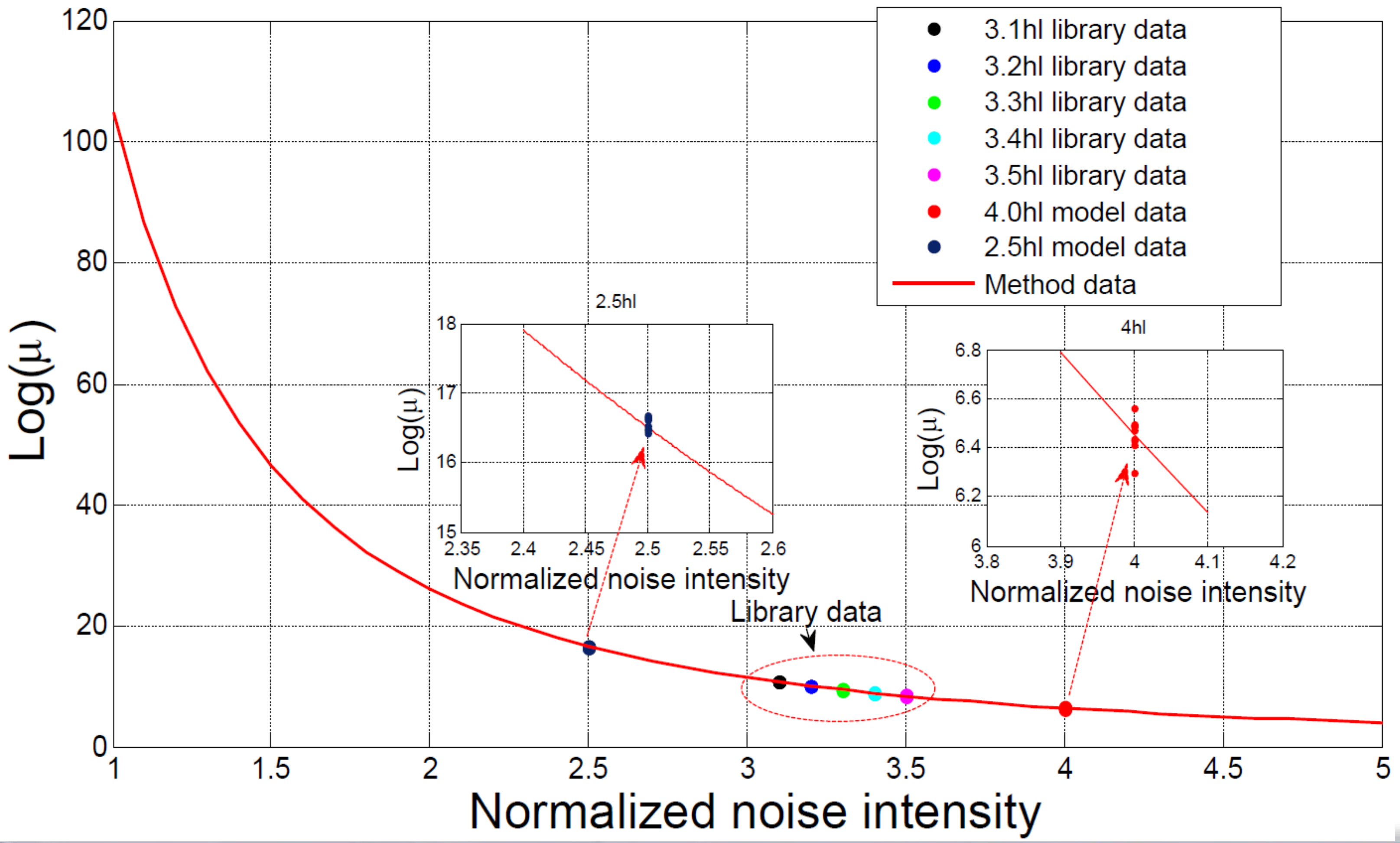


The **exponential distribution** may provide an excellent accuracy in modeling the probability distribution functions of a STT-RAM cell retention time.

$$P(t_{ret}) = \begin{cases} \frac{1}{\mu} e^{-\frac{1}{\mu} t_{ret}} & t_{ret} \geq 0, \\ 0 & t_{ret} < 0. \end{cases}$$

Step2: Transfer Statistical Information

How to determine the **statistical information of retention time** for any "**desired STT-RAM cell**"?

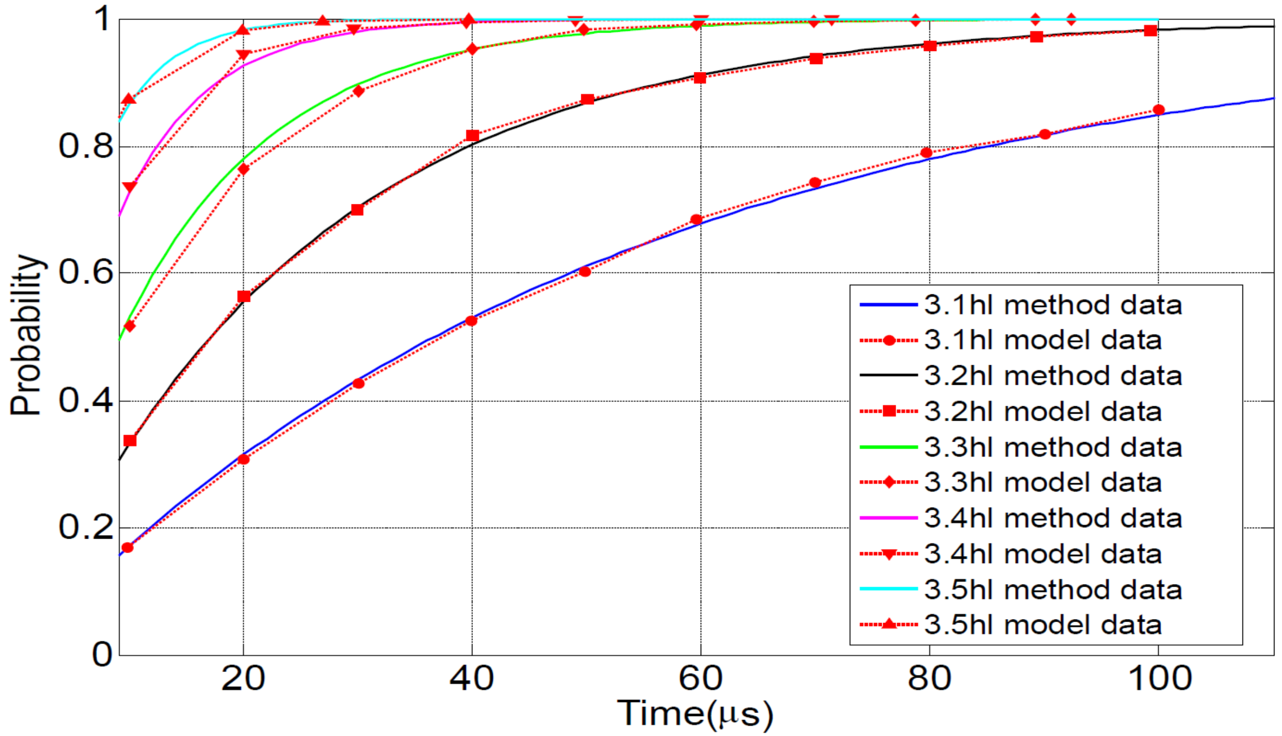


We propose to use a **noise intensity to retention time transfer function (N-R function)** to facilitate the associated information transfer:

$$log(\mu) = a/\beta^2 + b$$

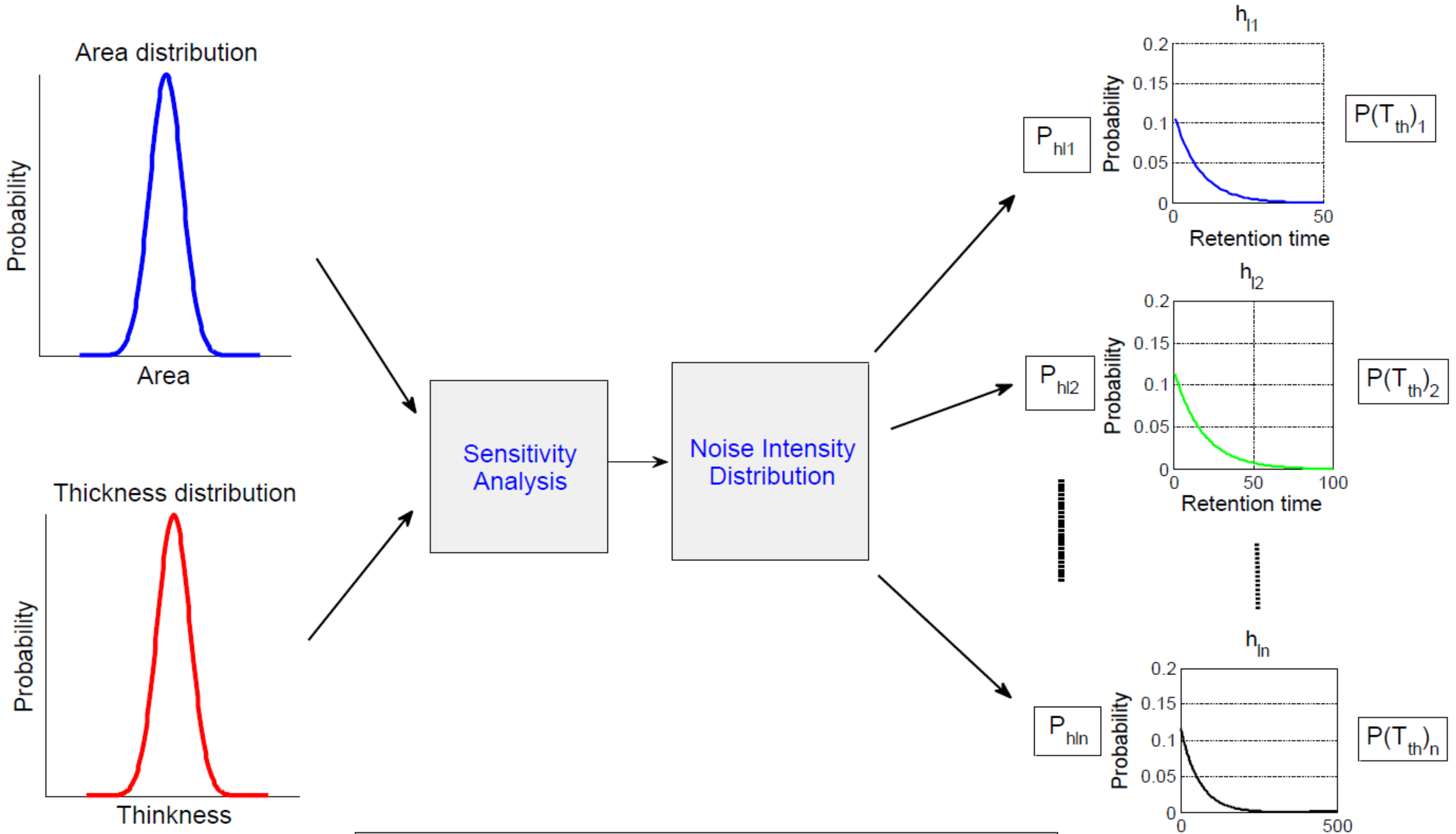
Step 3 and 4: Recover and calculate

The retention time **distribution recovery** for the "desired STT-RAM cell" can be easily conducted based on the **predicted mean** and **the exponential distribution function**.



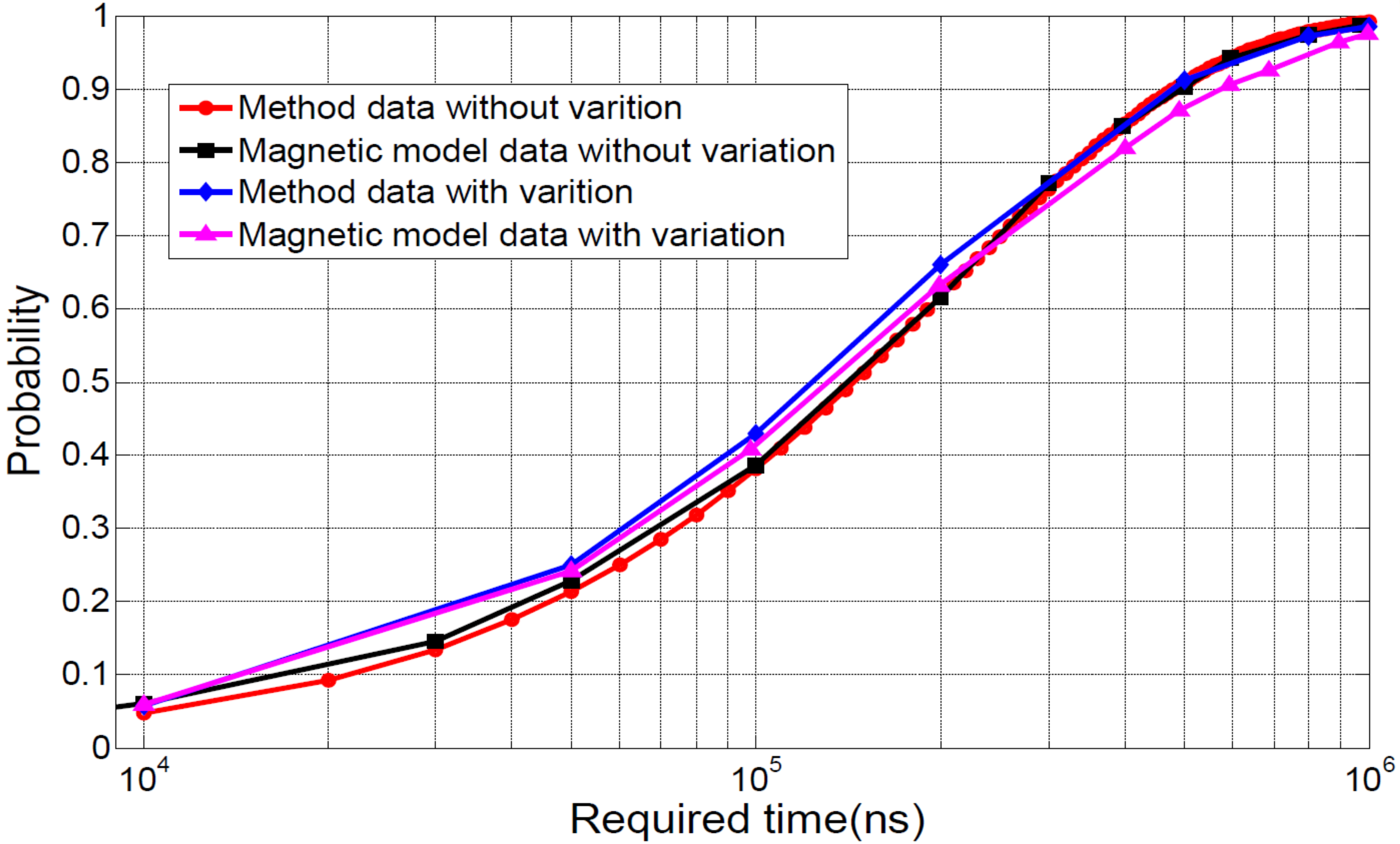
The **retention failure rate P_f** of "desired STT-RAM cell" can be easily captured by combining the **cumulative distribution functions (CDF) of retention time** and the **N-R function**.

Variation aware mode



$$P(T_{th})=P(T_{th})_1{*}P_{hl1}+P(T_{th})_2{*}P_{hl2}+\ldots P(T_{th})_n{*}P_{hln}$$

Variation aware mode results



We assume that both of MTJ **area** and **thickness** suffer from 2% **variations**.

The **noise intensity** is assumed to be 3.2h.

The results **with variations** are **very close** to those **without variations.**

Outline

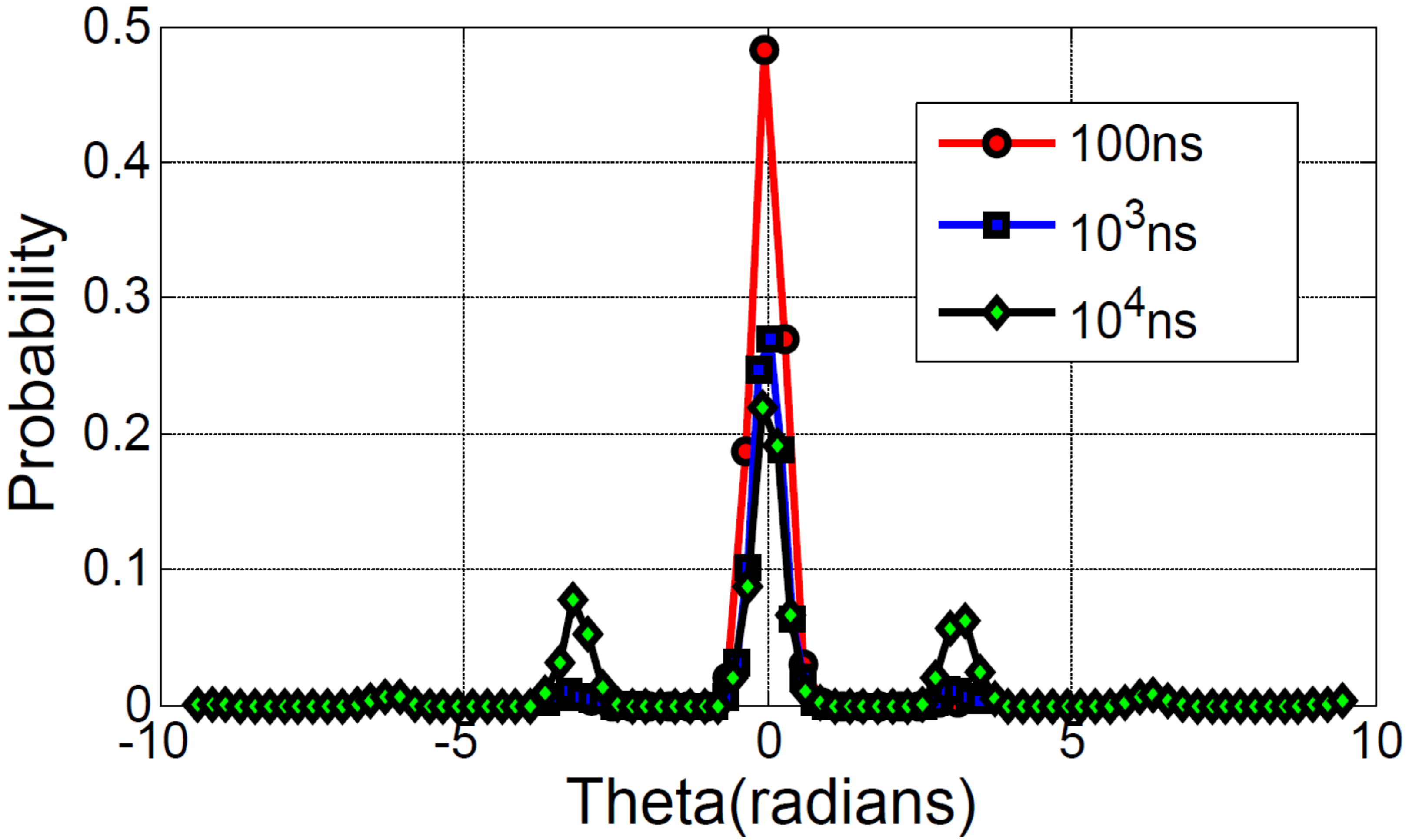
- Introduction
- Modeling and Validation
- Methodology and Simulation
- Results

--- Sensitivity Analysis of Initial Angle

--- Compare to traditional method

- conclusions

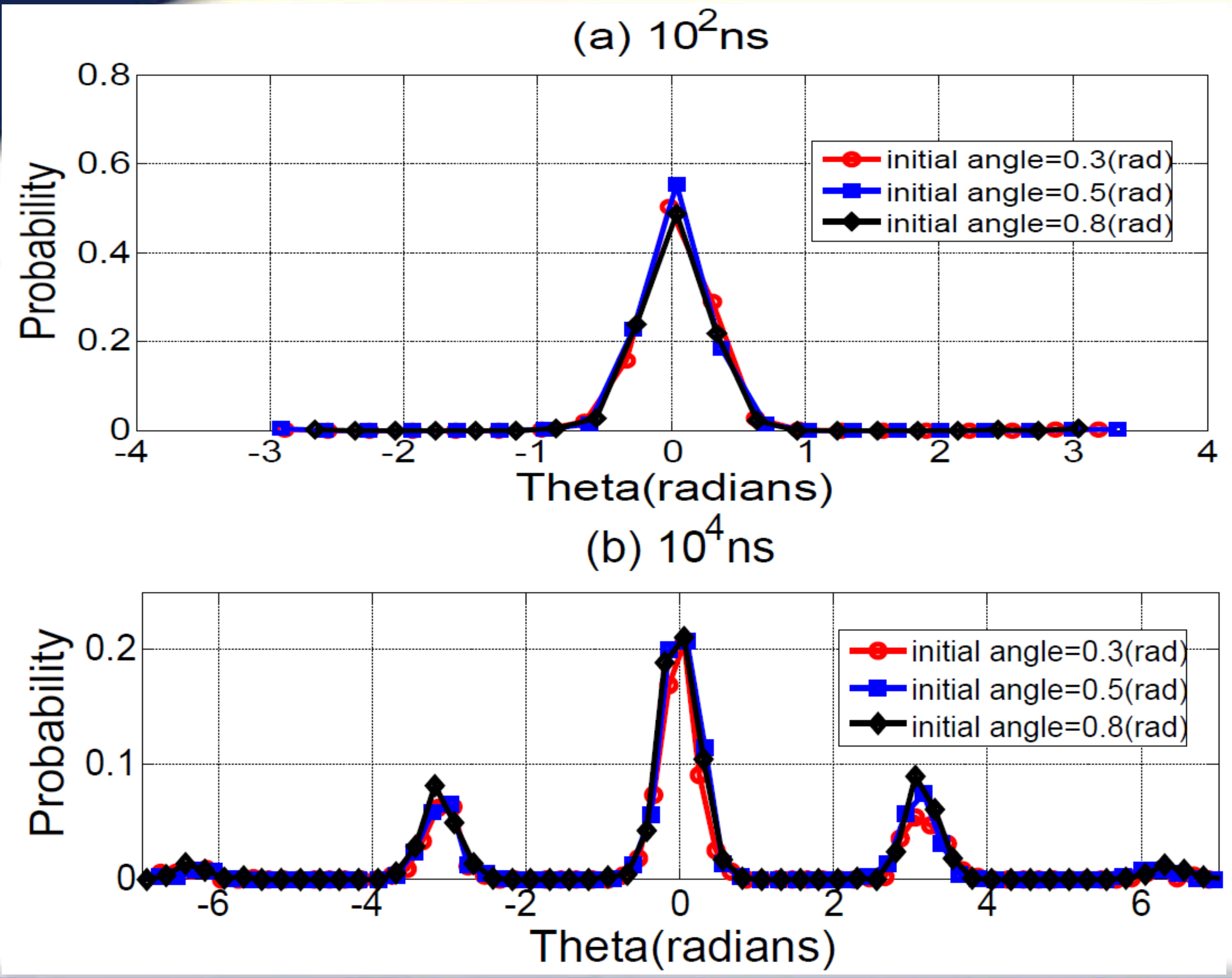
A dynamic changing of theta’s distribution



Observations:

- ❖ The **PDF of θ exhibits a very symmetric shape** around 0.
- ❖ As the **simulation time increases the bit flipping rate increase**.

Different initial angles



Our results clearly show that:

the initial angle does not impact the **retention time distribution**

The retention error rate comparison between our model and the traditional method

V scaling	P_t	noise intensity	P_m	T_p
$\Delta V/2^2$	0.1031	$2hl$	4×10^{-7}	$64ms$
$\Delta V/1.5^2$	5.5×10^{-5}	$1.5hl$	1.2×10^{-14}	$1s$
$\Delta V/3^2$	0.07	$3hl$	2.79×10^{-4}	$26.5us$

- **The traditional method always produces a pessimistic result.**
- **Previous works overestimate** the retention error rate and **adopt an unnecessarily high refresh frequency.**

Outline

- Introduction
- Modeling and Validation
- Methodology and Simulation
- Results
- conclusions

Conclusions

STT-RAM greatly suffers from a **unique data retention error** introduced purely by the **intrinsic thermal noise field**. It's very **difficult but crucial** to accurate calibration of such a type of error.

- ✓ We propose a retention time acceleration analysis method, namely "**Create-Transfer-Recover-Calculate**" to capture the **distribution** and **full statistical information** of retention time at interested time.
- ✓ Our method can **efficiently produce more precise results** than previous solutions, **providing a clear guidance for architectural designers**.
- ✓ Our result shows that the **initial angle has little impact on STT-RAM retention time**.
- ✓ At last we reveal that the retention time estimated by **previous works is too pessimistic**.

Thanks!

Q&A?